Plant Science 229 (2014) 181–192

Contents lists available at ScienceDirect

Plant Science

journal homepage: www.elsevier.com/locate/plantsci

The Wall-associated Kinase gene family in rice genomes

Luiz Felipe Valter de Oliveira[a,b,1], Ana Paula Christoff[a,b,1], Júlio Cesar de Lima[a,b], Bruno Comparsi Feijó de Ross[b], Gilberto Sachetto-Martins[c], Marcia Margis-Pinheiro[a], Rogerio Margis[a,b,d,*]

[a] Programa de Pós-Graduação em Genética e Biologia Molecular, Universidade Federal do Rio Grande do Sul, Brazil
[b] Centro de Biotecnologia e Programa de Pós-Graduação em Biologia Celular e Molecular, Universidade Federal do Rio Grande do Sul, Brazil
[c] Departamento de Genética, Universidade Federal do Rio de Janeiro, Brazil
[d] Departamento de Biofísica, Universidade Federal do Rio Grande do Sul, Brazil

ARTICLE INFO

Article history:
Received 19 July 2014
Received in revised form
10 September 2014
Accepted 12 September 2014
Available online 22 September 2014

Keywords:
WAK
Wall-associated Kinase
RLK
Rice
Oryza sativa

ABSTRACT

The environment is a dynamic system in which life forms adapt. Wall-Associated Kinases (*WAK*) are a subfamily of receptor-like kinases associated with the cell wall. These genes have been suggested as sensors of the extracellular environment and triggers of intracellular signals. They belong to the ePK superfamily with or without a conserved arginine before the catalytic subdomain VIB, which characterizes RD and non-RD *WAKs*. *WAK* is a large subfamily in rice. We performed an extensive comparison of WAK genes from *A. thaliana* (*AtWAK*), *O. sativa japonica* and *indica* subspecies (*OsWAK*). Phylogenetic studies and WAK domain characterization allowed for the identification of two distinct groups of WAK genes in *Arabidopsis* and rice. One group corresponds to a cluster containing only *OsWAKs* that most likely expanded after the *monocot–dicot* separation, which evolved into a non-RD kinase class. The other group comprises classical RD-kinases with both *AtWAK* and *OsWAK* representatives. Clusterization analysis using extracellular and kinase domains demonstrated putative functional redundancy for some genes, but also highlighted genes that could recognize similar extracellular stimuli and activate different cascades. The gene expression pattern of *WAKs* in response to cold suggests differences in the regulation of the *OsWAK* genes in the *indica* and *japonica* subspecies. Our results also confirm the hypothesis of functional diversification between *A. thaliana* and *O. sativa* WAK genes. Furthermore, we propose that plant *WAKs* constitute two evolutionarily related but independent subfamilies: *WAK*-RD and *WAK*-nonRD. Recognition of this structural division will further provide insights to understanding *WAK* functions and regulations.

1. Introduction

Plants use several environmental factors, such as light, temperature, gravity and water as signals for different physiological processes, including growth, flowering and dormancy. However, changes in the environment can perturb the metabolic homeostasis of plants, inducing abiotic stresses [1]. The ability to recognize signals from the environment and other adjacent cells and activate appropriate and specific downstream signaling cascades is important to plant responses. Until the late 80s, it was not clear how plant cells communicate with the environment. In 1990, the first receptor kinase protein (RLK – Receptor-Like Kinase) was described in plants [2]. Since then, several different plant *RLK* gene subfamilies have been identified, with more than 600 and 1100 genes in *Arabidopsis thaliana* and *Oryza sativa*, respectively [3–6].

The *RLK* family belongs to a protein kinase superfamily (ePK) that comprises all of the protein kinases identified in eukaryotes. The ePK superfamily plays an important role in post-translational protein modification through phosphorylation activity [7]. Members of the ePK superfamily possess 12 conserved subdomains. The subdomains I to IV are part of the amino-terminal lobe, associated with ATP binding. Subdomain V corresponds to a link between the amino and the carboxyl-terminal lobe. The subdomains VIA, VIB, VII, VIII, IX, X and XI are part of the carboxyl-terminal lobe, which is associated with peptide-substrate binding and phosphorylation activity [7–10]. The subdomain VIB contains the catalytic motif DxxxxN, which directly participates in substrate phosphorylation, and the presence of a conserved arginine (R) residue preceding the

* Corresponding author at: Centro de Biotecnologia, sala 213, prédio 43431, Universidade Federal do Rio Grande do Sul – UFRGS, P.O. Box 15005, CEP 91501-970 Porto Alegre, RS, Brazil. Tel.: +55 51 3308 7766; fax: +55 51 3308 6072.
E-mail addresses: rogerio.margis@ufrgs.br, rogerio.margis@gmail.com (R. Margis).
[1] Theses authors contributed equally to this manuscript.

http://dx.doi.org/10.1016/j.plantsci.2014.09.007
0168-9452/

DxxxxN motif classifies these kinases as RD-kinases. Conversely, kinases without this R residue are referred to as non-RD kinases. Additionally, the subdomain VIII is part of the activation segment along with the conserved APE residues [11]. Several mechanisms regulate the activities of these two classes, which are involved in different biological processes [10,12,13].

The RLK family is involved in a wide range of processes in plants, such as hormone signaling [14–16], tissue development [17–19], reproduction [20–23], symbiosis [24], nodulation [25], biotic and abiotic stress responses [15,26–28]. A comparative analysis of RLK genes in animals and plants suggested a large gene expansion restricted to plants that most likely occurred after the divergence of these two kingdoms [3]. In general, RLK proteins have an extracellular domain (EC) and a kinase domain. However, it is also possible to find two other protein structural arrangements associated with RLKs: (i) genes that only encode a kinase domain (RLCK – Receptor-Like Cytoplasmatic Kinase) [3,6,29,30] and (ii) genes that only encode the extracellular domain (RLP – Receptor-Like Protein) [29].

In vascular plants there is a subfamily of RLKs known as *WAKs* (Wall-Associated Kinases) [31]. *WAKs* were initially described in *Arabidopsis thaliana* as a cluster of five (*WAK1-5*) [32], and 22 (WAKL1-WAKL22) genes [33]. The *WAK* family in *A. thaliana* is expressed in diverse tissue-specific and developmental patterns [32,34]. Currently, *WAK*s are the only known proteins that physically link cell walls to the plasma membrane, directly transmitting extracellular signals from their EC domain to the cytoplasm, through the cytoplasmic kinase domain [31,34–36]. A previous study of *O. sativa* genome identified 125 genes belonging to the *WAK* subfamily, *OsWAKs* [29], and classified these genes according to the presence/absence of EC and kinase domains. In this work, we extend the current phylogenetic knowledge of the *WAK* subfamily among *A. thaliana*, *O. sativa indica* and *japonica* subspecies to characterize the residues of the kinase domain. Furthermore, we identify potential redundant proteins, through clustering EC and kinase domains, and characterize the expression of twenty-two *indica* and *japonica Os*WAK genes in response to cold treatment. Our results suggest the need for a novel classification of the WAK subfamilies: WAK-RD and WAK-nonRD, where WAK-nonRDs are restricted to monocots. These findings will be important to drive further functional studies concerning the *WAK* subfamily.

2. Materials and methods

2.1. OsWAK identification, gene prediction, domain characterization and re-annotation

OsWAK searches were performed using all genes previously described for this subfamily in *O. sativa* ssp *japonica* in the *Rice Genome Annotation Project Version 6.1* [http://rice.plantbiology.msu.edu/]. We used tBLASTx [37] from the Gramene database (http://www.gramene.org) against *O. sativa* ssp *indica* genome (Supplementary Fig. 1). *OsWAK japonica* genomic sequence was also used as a query to find misannotated genes. The lack of predicted protein domains could undermine this search. To re-annotate the *japonica OsWAK* genes, we performed protein domain analysis through SMART (Simple Modular Architecture Research Tool) [38] for all annotated splicing for all *japonica OsWAK*. Afterward, we used genes identified in the *indica* genome with different protein features than observed for *japonica OsWAKs* that were previously annotated or from our first tBLASTx procedure. This procedure was used to determine the potential for alternative splicing of *japonica OsWAK*. We also performed splicing prediction using GENSCAN for the *OsWAK* genomic sequences followed by protein domain analysis through SMART for all predicted CDS. All hits confirmed as *WAK* genes from the *japonica* genome were used in a new round of tBLASTx in the *indica* genome.

The sequences recovered from the *indica* genome were subjected to gene splicing prediction through GENSCAN [39] using the *A. thaliana* matrix. All proteins deduced from predicted CDS were analyzed with SMART and Pfam [40] to define their protein domains. Sequences that presented a domain related to a kinase and/or EC in the predicted protein were selected as putative *OsWAK* genes. Subsequently, we performed BLASTp for all putative *indica OsWAK* genes against the Rice *japonica* pseudomolecules protein database [ftp://ftp.plantbiology.msu.edu/pub/data/EukaryoticProjects/o_sativa/annotation_dbs/pseudomolecules/version_6.1/all.dir/all.pep]. Finally, the best hit for the putative *indica OsWAK* genes was annotated as *japonica* ortholog and named as *indica OsWAK* (*OsWAK_ind*).

2.2. Multiple sequence alignment and domain characterization

The predicted protein sequences were used to perform a multiple sequence alignment with MUSCLE software [41]. The alignment was manually inspected and the *indels* individually adjusted from the multiple alignments matrix. From the resulting alignment, a graphical consensus of the kinase domain was performed for all *OsWAKs* and *AtWAK* using the WebLOGO 3.0 tool [42] with the consensus probability as the unit of the figure scale.

2.3. Clustering analysis

Similarity clustering analysis was performed using Circoletto [43], a pipeline in Perl to create a similarity visualization through Circos [44] using results from BLASTp. The *E*-value of 1E-90 was set as the parameter to create the edge information. To perform BLASTp, both EC domain data and kinase data were prepared independently from the same original sequences. The similarity sequence network was created independently for each dataset using the presence of local alignments from BLASTp as information of interaction among *WAKs*. Both networks were imported to Cytoscape [45] and the plug-in GraphMerge was used to perform a comparison between the two networks and to recover proteins with homologies in both EC and kinase domains or sequences with similarities only in the EC domain.

2.4. Phylogenetic analysis

The phylogenetic analysis was performed using 95 OsWAK protein sequences from *O. sativa japonica*, 85 OsWAK proteins from *indica* and 24 AtWAK proteins from *A. thaliana*. In addition, one protein from the RLK family with a corresponding orthologous gene present in both *O. sativa* (Os01g02560) and *A. thaliana* (AT1G67000) species, which did not belong to the WAK subfamily, was used as an out group. Four independent phylogenetic approaches were applied to the WAK amino acid sequences alignment, Bayesian, Maximum Likelihood, Neighbor-Joining and Maximum Parsimony. All phylogenies were performed using the same protein alignment. The Bayesian phylogenetic analysis was performed using MrBayes [46] setting the model as “mixed” and 15,000,000 generations. The Maximum Likelihood phylogeny was performed using PhyML [47]. The MEGA4 software [48] was employed for Neighbor-Joining and Maximum Parsimony, using 1000 bootstrap replicates. For the Neighbor-Joining phylogeny, the poison-correction substitution model was used.

2.5. Plant material, cold stress and RT-qPCR analysis

Cold stress experiments were conducted using *O. sativa* spp *japonica* cv. nipponbare and *O. sativa* spp *indica* cv. Embrapa Taim